International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 02, 2020
ISSN: 1475-7192

A Survey on Fake News Detection

[1] Sireddy.Naveen Kumar Reddy, [2]Dr.M.Sujatha

Abstract-- *Social media is double edged sword for news consumption on one side it is easy to access, low cost and other side fake news will spread widely, includes false information. The wide spread of this fake news results in negative impact on society and individuals. Fake news is mainly created to misguide readers in order to believe information that is not true. Source, headline, body text and image or video are the content attributes for news. Source is the news article publisher. The short text that makes readers attention and describes the important topic in that article is called headline. The entire content about that article is present in body text and includes images or videos that are related to that article. Based on this attributes fake news characteristics are extracted*.

Keywords—*False news, training the data's.*

I INTRODUCTION

False news is the false news that is published to damage the name of organization, person, agency etc., this fake news effect to justice, public trust and democracy is increasing. So, it is important to detect fake news. The study of fake news mainly concentrated of four important things i) the untruth news it carries ii) patterns of propagation iii) style of writing iv) news spreaders and creators. Fake news is a threat to journalism, democracy and freedom of speech. Now a day's people are addicted to social media like facebook, twitter. So, fake news creators can easily create and publish news online faster when compared to traditional media television and newspaper. Social media provides features to forward, review, share and encourages the users to discuss about the news in online. All this features of social media encourages fake news creators to create news, publish it and spread that fake news. Some fundamental theorizers are used for analysis of fake news I) style based fake news analysis ii) propagation based fake news analysis iii) user based fake news analysis. Analyzing or detecting fake news is called fact checking. Fact checking is of two types i) manual fact checking ii) automatic fact checking. Manual fact checking is divided into expert- based fact checking and crowd -sourced fact checking. Expert based fact checking is easy to process and gives accurate results but it is costly. Crowed-sourced fact checking is difficult when compared to expert based fact checking. Automatic fact checking involves two steps fact extraction and fact checking. In fact extraction information is extracted from online websites and documents. In the process of fact checking is done by comparing the extracted news with facts stored.

[1] *Department of Electronics and Communication, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai-602105, Tamil Nadu, India, Email:sireddynaveen@gmail.com*

[2] *Department of Electronics and Communication, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Chennai-602105, Tamil Nadu, India, Email:sujathamk.sse@saveetha.com*

Received: 23 Dec 2019 | Revised: 05 Jan 2020 | Accepted: 27 Feb 2020

II LITERATURE SURVEY

1) Akshay jain et., al., proposed a method to detect fake news and explained how to implement it on facebook, which is the most popular social media platform s. naïve bayel classification model is used in this method and this will identify whether a news is fake or real. According to naïve bayes classifier all the features are not related to each other. A feature presence or absence does not influence the absence or presence of another feature. To obtain such kind of classifiers one single algorithm is not enough. So, a family of algorithms that are on common principle used to made these classifiers. The algorithms identify number of times a particular event has occurred. On implementing bayes theorem it is also possible to calculate probability of news. Future work: more data can be used for training purpose porter stemming method of deleting the common morphological ending more words in English.

2) shivam et., al., in research community detection of fake news has become a popular topic since in modern days journalism is at its peak. This paper gives information about different types of news and its impact on readers. Fake news is defined as news that is not true and is created to convince the readers to believe in something that is not true. The author explains with an example that is fake news did Palestinians recognizes texas as part of mexico. After publishing this news in just 4 days it got 1.5 million likes. This can make us to understand how this story is trending on social media. This implies that readers focus on headline and multimedia content of the story. In general a news has 3 things they are i) headline ii) multimedia iii) body that is the content. In a study it has been proven that 70% of social media users just read the headline and they comment or share it.

3) Terry traylor et., al., explained about research process, technical analysis, technical linguistics work, performance of a classifier and results. Fake news detection is possible in many ways. In fake news detection algorithm every document present in the document collection, the document is paragraphs are tokenized and counted. In that each paragraph is checked for quotes. The negative attributions receives a -1 score and positive attribution receives a+1 score. Then these positive attributions negative attributions are added and are named as A- score. If the A-score is greater is greater than zero, then that document is real. The document is fake when its a-score is lesser than 0. In this paper it was clearly explained fake news detection for large documents with one extraction feature using machine learning.

4)Kyeong-hwan et., al.,To verify whether a news is fact or not a method named bidirectional multiperspective matching for natural language sentence (BiMPM) is used. These methods can perform well in task of matching sentence. But this (BiMPM) has certain limitations. Due these limitations, the performance of this BiMPM is lower if the input sentence length is longer. So, that it is difficult for accurate judgment. To overcome this limitations article abstraction a new sentence matching technique is introduced. By these two techniques the performance of fake news detection will increase. In this process the system first receives the input using related article finder modules from the fact DB the system searches related articles for the input. Then through article abstraction module previously searched articles are abstracted into sentence. This is received by BiMPM and entity setting modules. These both will calculate answers accordingly. Then answer calculator will tell whether the news is fact or not by comparing answers with matching modules. The BiMPM considers new words as noise and it cannot identify

meanings and differences between one or two words. So, by adding article abstraction it is possible to overcome these disadvantages and helps to improve performance.

5)Rohit kumar et., al., News obtaining from social media has advantages and disadvantages. Advantages are consumes less time, user friendly, can access easily, socially relevant news are easily conveyable and every minute updates. Disadvantage is that news can be manipulated easily by other sites according to once own interest and personal opinions. Manipulation of news intentionally to damage name of agency, organization and a person is called fake news. To classify data sets machine learning, language processing and deep learning methods are used. Jupyter notebook platform is used. Jupyter notebook is a open software which can execute science, codes, contents, It is possible to upload python files and codes and output will be in the form of matrix. To get high accuracy models of machine learning and deep learning techniques are used.

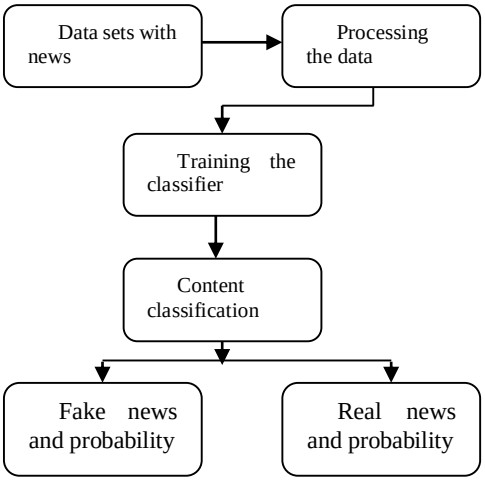
6)Syed ishfaq et., al., classification of information is possible with some limitations using machine learning. This paper explains detection of fake news using various machine learning methods. In general there are five types of fake news. They are i)user friendly ii)visual based iii)knowledge based iv) style -based v) stance based user based fake news is created by fake accounts which represents gender, political affiliations, culture visual based fake news is a graphical post that contains morphed image and videos. Knowledge based fake news is about some scientific unresolved issues explanations. Style based news are news that is copied from some accredited journals by some journalists. Stance-based is a statements of truth but it changes meaning and purpose of news. Computing hierarchical features is the main characteristic feature of deep learning. Deep learning methods are deep boltzmann machine, convolutional neural network, deep autoencoder model, deep neural network.

7)bhavika bhutani et., al., explained about fake news detection using sentiment analysis. People are addicted to social media because news and articles are available with just a click. In this method of detecting fake news sentiment plays a important role to improve accuracy and the results are will when compared to other methods. From datasets like politifact, emergent merged data sets are prepared. The data will undergo different pre-processing techniques such as count vectorizer, document frequency vectorizer, bigrams, trigrams. In this process they used ti-idf vectorizer with cosine to build vocabulary. After this process in order to predict the sentiment of news statements of tested data naives bayes classifier is used. Then in that merged data additional columns like tf-idf scores, cosine and sentiments scores are added. Naive bayes and random forest are used to build training model. Finally performance is evaluated and accuracy is compared. With n-grams and without n-grams count vectorizer is also applied. Many people are consuming news from social media rather than traditional media. This paper explains about different techniques of text preprocessing. In future application in other neural networks can be done.

8)Namwon kim et., al., explained a unified key sentence information method for fake news detection. To make a key sentence from article unified word vector is used and this is made by extracting article and merging word vector for every key sentence. Sentence matching is a technique of checking whether two sentences are same or not. It is also called as natural language processing. Other techniques named recurrent neural network is used to find meaning of sentences. It can process large amount of data. Limitations of this recurrent neural network can be overcome by long short term memory. If a sentence questions A and a set of sentences in a article B is given. Then our model matches A and B to find out whether A is true or not. It consists of 5 layers. i) key sentence ii) word

representation layer iii) context representation layer iv) matching representation layer v) decision layer. In key sentence layer question is decomposed and article into words. If we compare this model with previous BiMPM model the accuracy is increased from 64% to 69%.

III PROPOSED APPROACH



The first step in the classification is collecting the data sets from LIAR or Buzzfeed. The model is trained with the data sets. Then the processed go through the classifier like Naïve-Bayes classifier, logistic regression classifier, support vector machine classifier, stochastic gradient descent classifier and Random forest classifier. All these machine learning algorithms and natural language processing algorithms are used to classify fake news articles using libraries from anaconda.

After setting the entire required environment, the classification of news can be done. The results that are obtained in the module are probability of truth and whether the news is fake or real.

IV RESULTS

Table 1: Summary of probability of truth based on input news

S.No	Input news	Fake or real	Probability of truth
1	A flight from Atlanta to Houston was cancelled due to terrorist dry run	fake	0.485419
2	Bill McCollum has "recovered $200 million in Medicaid fraud."	real	0.5659615
3	When Obama was sworn into office, he DID NOT use the Holy Bible, but instead the Kuran (Their equivalency to our Bible, but	fake	0.4757452

Received: 23 Dec 2019 | Revised: 05 Jan 2020 | Accepted: 27 Feb 2020

	very different beliefs).		
4	Chemical weapons have been used probably 20 times since the Persian Gulf War	real	0.6322886

V CONCLUSION

Thus a machine learning model can be developed for checking the truth of the news. This model is a combination of different machine learning algorithms. This model can also produce the probability of the truth of the news.

REFERENCES

1. Jain, Akshay, and Amey Kasbe. "Fake news detection." In 2018 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS), pp. 1-5. IEEE, 2018..
2. Parikh, Shivam B., and Pradeep K. Atrey. "Media-rich fake news detection: A survey." In 2018 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), pp. 436-441. IEEE, 2018.
3. Traylor, Terry, Jeremy Straub, and Nicholas Snell. "Classifying fake news articles using natural language processing to identify in-article attribution as a supervised learning estimator." In 2019 IEEE 13th International Conference on Semantic Computing (ICSC), pp. 445-449. IEEE, 2019.
4. Kim, Kyeong-Hwan, and Chang-Sung Jeong. "Fake News Detection System using Article Abstraction." In 2019 16th International Joint Conference on Computer Science and Software Engineering (JCSSE), pp. 209-212. IEEE, 2019.
5. Kaliyar, Rohit Kumar. "Fake news detection using a deep neural network." In 2018 4th International Conference on Computing Communication and Automation (ICCCA), pp. 1-7. IEEE, 2018..
6. Manzoor, Syed Ishfaq, and Jimmy Singla. "Fake News Detection Using Machine Learning approaches: A systematic Review." In 2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI), pp. 230-234. IEEE, 2019.
7. Bhutani, Bhavika, Neha Rastogi, Priyanshu Sehgal, and Archana Purwar. "Fake news detection using sentiment analysis." In 2019 Twelfth International Conference on Contemporary Computing (IC3), pp. 1-5. IEEE, 2019.
8. Kim, Namwon, Deokjin Seo, and Chang-Sung Jeong. "FAMOUS: Fake News Detection Model Based on Unified Key Sentence Information." In 2018 IEEE 9th International Conference on Software Engineering and Service Science (ICSESS), pp. 617-620. IEEE, 2018.
9. Pandey, Amit, and Gyan Prakash. "Deduplication with Attribute Based Encryption in E-Health Care Systems." International Journal of MC Square Scientific Research 11, no. 4 (2019): 16-24.
10. Shahada, Shareefa Ahmad Abu, Suzan Mohammed Hreiji, and Shermin Shamsudheen. "IOT BASED GARBAGE CLEARANCE ALERT SYSTEM WITH GPS LOCATION USING ARDUINO." International Journal of MC Square Scientific Research 11, no. 1 (2019): 1-8.

Received: 23 Dec 2019 | Revised: 05 Jan 2020 | Accepted: 27 Feb 2020